



SEQUENCE LISTING

RECEIVED
OCT 26 2000
TECH CENTER 1600/2900
RECEIVED

OCT 30 2000

(1) GENERAL INFORMATION:

(i) APPLICANT: von Schaewen, Antje Dr. rer. nat.

(ii) TITLE OF INVENTION: Plant gntI sequences and the use thereof for
the production of plants having reduced or lacking N-acetyl
glucosaminyl transferase I (GnTI) activity

TECH CENTER 1600/2900

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: von Schaewen, Antje Dr. rer. nat.
- (B) STREET: Natruperstrasse 169a
- (C) CITY: Osnabrueck
- (D) COUNTRY: Germany
- (E) ZIP: D-49076

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA

- (A) APPLICATION NUMBER: 09/591,466
- (B) FILING DATE: 06/09/2000
- (C) CLASSIFICATION: 536

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1669 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Solanum tuberosum
- (B) STRAIN: Desiree
- (D) DEVELOPMENTAL STAGE: Sink organ
- (F) TISSUE TYPE: Mesophyll
- (G) CELL TYPE: Leaf cells

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Lambda ZAP II (Eco RI)
- (B) CLONE: gntI-A1(K)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 659..667
- (D) OTHER INFORMATION:/function= "Asn codon in this
context is a potential glycosylation site"
/product= "N-glycosylation consensus sequence"

```

1      /phenotype= "N-glycans modulate protein
2      properties"
3      /standard_name= "N-glycosylation site"
4      /label= pot-CHO
5      /note= "GnTI-coding sequences from animals do not
6      contain this feature"
7
8      (ix) FEATURE:
9          (A) NAME/KEY: CDS
10         (B) LOCATION:53..1393
11         (C) IDENTIFICATION METHOD: experimental
12         (D) OTHER INFORMATION:/codon_start= 53
13             /function= "initiates complex N-glycans on
14             secretory glycoproteins"
15             /EC_number= 2.4.1.101
16             /product=
17             "beta-1,2-N-acetylglucosaminyltransferase I"
18             /evidence= EXPERIMENTAL
19             /gene= "cgl"
20             /standard_name= "gntI"
21             /label= ORF
22             /note= "first gntI sequence from potato (unpublished)"
23
24      (ix) FEATURE:
25          (A) NAME/KEY: 5'UTR
26          (B) LOCATION:15..52
27
28      (ix) FEATURE:
29          (A) NAME/KEY: 3'UTR
30          (B) LOCATION:1394..1655
31
32      (ix) FEATURE:
33          (A) NAME/KEY: CDS
34          (B) LOCATION:80..139
35          (D) OTHER INFORMATION:/function= "membrane anchor (amino
36          acids 10-29)"
37          /product= "hydrophobic amino acid stretch in GnTI"
38          /standard_name= "membrane anchor of a type II
39          Golgi protein"
40          /note= "identified by comparison with GnTI sequences
41          from animals"
42
43      (ix) FEATURE:
44          (A) NAME/KEY: misc_feature
45          (B) LOCATION:1..14
46          (D) OTHER INFORMATION:/function= "used for cloning the
47          cDNA library in Lambda ZAPII"
48          /product= "EcoRI/NotI-cDNA adapter"
49          /number= 1
50
51      (ix) FEATURE:
52          (A) NAME/KEY: misc_feature
53          (B) LOCATION:1656..1669
54          (D) OTHER INFORMATION:/product= "EcoRI/NotI-cDNA adapter"
55          /number= 2
56
57
58      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
59
60      GAATTCGCGG CCGCCTGAGA AACCCCTCGAA TTCAATTTCG CATTGGCAG AG ATG      55
61                                         Met
62                                         1
63

```

1	AGA GGG AAC AAG TTT TGC TTT GAT TTA CGG TAC CTT CTC GTC GTG GCT	103
2	Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val Ala	
3	5 10 15	
4	GCT CTC GCC TTC ATC TAC ATA CAG ATG CGG CTT TTC GCG ACA CAG TCA	151
5	Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln Ser	
6	20 25 30	
7	GAA TAT GTA GAC CGC CTT GCT GCA ATT GAA GCA GAA AAT CAT TGT	199
8	Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His Cys	
9	35 40 45	
10	ACA AGT CAG ACC AGA TTG CTT ATT GAC AAG ATT AGC CAG CAA GGA	247
11	Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln Gly	
12	50 55 60 65	
13	AGA GTA GTA GCT CTT GAA GAA ATG AAG CAT CAG GAC CAG GAG TGC	295
14	Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu Cys	
15	70 75 80	
16	CGG CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG GGC ATA AAA	343
17	Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile Lys	
18	85 90 95	
19	AAG TTA ATC GGA GAT GTG CAG ATG CCA GTG GCA GCT GTA GTT GTT ATG	391
20	Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val Met	
21	100 105 110	
22	GCT TGC AGT CGT ACT GAC TAC CTG GAG AGG ACT ATT AAA TCC ATC TTA	439
23	Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile Leu	
24	115 120 125	
25	AAA TAC CAA ACA TCT GTT GCA TCA AAA TAT CCT CTT TTC ATA TCC CAG	487
26	Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln	
27	130 135 140 145	
28	GAT GGA TCA AAT CCT GAT GTA AGA AAG CTT GCT TTG AGC TAT GGT CAG	535
29	Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly Gln	
30	150 155 160	
31	CTG ACG TAT ATG CAG CAC TTG GAT TAT GAA CCT GTG CAT ACT GAA AGA	583
32	Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu Arg	
33	165 170 175	
34	CCA GGG GAA CTG GTT GCA TAC TAC AAG ATT GCA CGT CAT TAC AAG TGG	631
35	Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp	
36	180 185 190	
37	GCA TTG GAT CAG CTG TTT CAC AAG CAT AAT TTT AGC CGT GTT ATC ATA	679
38	Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile Ile	
39	195 200 205	
40	CTA GAA GAT GAT ATG GAA ATT GCT GCT GAT TTT TTT GAC TAT TTT GAG	727
41	Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe Glu	
42	210 215 220 225	
43	GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG GCT ATT TCT	775
44	Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile Ser	
45	230 235 240	
46	TCT TGG AAT GAC AAT GGA CAA AGG CAG TTC GTC CAA GAT CCT GAT GCT	823
47	Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp Ala	
48	245 250 255	

1	CTT TAC CGC TCA GAC TTT TTT CCT GGT CTT GGA TGG ATG CTT TCA AAA	871
2	Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser Lys	
3	260 265 270	
4		
5	TCA ACT TGG TCC GAA CTA TCT CCA AAG TGG CCA AAG GCT TAC TGG GAT	919
6	Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp	
7	275 280 285	
8		
9	GAC TGG CTA AGG CTG AAA GAA AAT CAC AGA GGT CGA CAA TTT ATT CGC	967
10	Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile Arg	
11	290 295 300 305	
12		
13	CCA GAA GTT TGC AGA ACG TAC AAT TTT GGT GAG CAT GGT TCT AGT TTG	1015
14	Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu	
15	310 315 320	
16		
17	GGG CAG TTT AAG CAG TAT CTT GAG CCA ATT AAG CTA AAT GAT GTC	1063
18	Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val	
19	325 330 335	
20		
21	CAG GTT GAT TGG AAG TCA ATG GAC CTA AGT TAC CTT TTG GAG GAC AAC	1111
22	Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp Asn	
23	340 345 350	
24		
25	TAT GTG AAA CAC TTT GGC GAC TTG GTT AAA AAG GCT AAG CCC ATC CAC	1159
26	Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile His	
27	355 360 365	
28		
29	GGA GCT GAT GCT GTT TTG AAA GCA TTT AAC ATA GAT GGT GAT GTG CGT	1207
30	Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val Arg	
31	370 375 380 385	
32		
33	ATT CAG TAC AGA GAC CAA CTA GAC TTT GAA GAT ATC GCT CGA CAG TTT	1255
34	Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln Phe	
35	390 395 400	
36		
37	GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGG GCA GCA TAT AAA	1303
38	Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr Lys	
39	405 410 415	
40		
41	GGG ATA GTA GTT TTC CGG TTT CAA ACA TCT AGA CGT GTG TTC CTT GTT	1351
42	Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu Val	
43	420 425 430	
44		
45	TCC CCT GAT TCT CTT CGA CAA CTT GGA GTT GAA GAT ACT TAG	1393
46	Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr *	
47	435 440 445	
48		
49	CGAAGATATG ATTGGAGCCT GAGCAACAAT TTAGACTTAT TTGGTAGGAT ACATTTGAAA	1453
50		
51	GAGCTGACAC GAAAAGTATG ACTACCAGTA GCTACATGCA ACATTTAAC GTTAATGGAA	1513
52		
53	GGAACCCACT GCTTATTGTT GGAATGGATG AATCATCAC ACCATCCTATT ATTCAAGTTT	1573
54		
55	ACAAACATAA AGAGGAAATG TTGCCCTATA AAAACAAATT TTTTGTTC T AAGAAGGAAC	1633
56		
57	GTTACGATTA TGAGCAACTT TGGCGGCCGC GAATTC	1669
58		
59	(2) INFORMATION FOR SEQ ID NO: 2:	
60		
61		
62	(i) SEQUENCE CHARACTERISTICS:	
63	(A) LENGTH: 446 amino acids	

```

1 (B) TYPE: amino acid
2 (D) TOPOLOGY: linear
3
4 (ii) MOLECULE TYPE: protein
5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
6
7 Met Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val
8 1 5 10 15
9
10 Ala Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln
11 20 25 30
12
13 Ser Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His
14 35 40 45
15
16 Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln
17 50 55 60
18
19 Gly Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu
20 65 70 75 80
21
22 Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
23 85 90 95
24
25 Lys Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val
26 100 105 110
27
28 Met Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile
29 115 120 125
30
31 Leu Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
32 130 135 140
33
34 Gln Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly
35 145 150 155 160
36
37 Gln Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu
38 165 170 175
39
40 Arg Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
41 180 185 190
42
43 Trp Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile
44 195 200 205
45
46 Ile Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe
47 210 215 220
48
49 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
50 225 230 235 240
51
52 Ser Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp
53 245 250 255
54
55 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
56 260 265 270
57
58 Lys Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
59 275 280 285
60
61 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
62 290 295 300
63

```

1 Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
 2 305 310 315 320
 3 Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
 4 325 330 335
 5 Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
 6 340 345 350
 7 Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
 8 355 360 365
 9 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
 10 370 375 380
 11 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln
 12 385 390 395 400
 13 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
 14 405 410 415
 15 Lys Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu
 16 420 425 430
 17 Val Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr
 18 435 440 445
 19 (2) INFORMATION FOR SEQ ID NO: 3:
 20
 21 (i) SEQUENCE CHARACTERISTICS:
 22 (A) LENGTH: 1737 base pairs
 23 (B) TYPE: nucleic acid
 24 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear
 26
 27 (ii) MOLECULE TYPE: cDNA to mRNA
 28
 29 (iii) HYPOTHETICAL: NO
 30
 31 (iv) ANTI-SENSE: NO
 32
 33 (vi) ORIGINAL SOURCE:
 34 (A) ORGANISM: Nicotiana tabacum
 35 (B) STRAIN: Samsun NN
 36 (D) DEVELOPMENTAL STAGE: Sink organ
 37 (F) TISSUE TYPE: Mesophyll
 38 (G) CELL TYPE: Leaf cells
 39
 40 (vii) IMMEDIATE SOURCE:
 41 (A) LIBRARY: Lambda ZAP II (Eco RI)
 42 (B) CLONE: gntI-A9(T)
 43
 44 (ix) FEATURE:
 45 (A) NAME/KEY: misc_feature
 46 (B) LOCATION: 733..741
 47 (D) OTHER INFORMATION:/function= "Asn codon in this
 48 context is a potential glycosylation site"
 49 /product= "N-glycosylation consensus sequence"
 50 /phenotype= "N-glycans modulate protein
 51 properties"
 52 /standard_name= "N-glycosylation site"
 53 /label= pot-CHO
 54 /note= "GnTI sequences from animals do not contain this

1	ACA CAG TCA GAA TAT GCA GAT CGC CTT GCT GCA ATT GAA GCA GAA	264
2	Thr Gln Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ile Glu Ala Glu	
3	480 485 490	
5	AAT CAT TGT ACA AGC CAG ACC AGA TTG CTT ATT GAC CAG ATT AGC CTG	312
6	Asn His Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu	
7	495 500 505	
9	CAG CAA GGA AGA ATA GTT GCT CTT GAA GAA CAA ATG AAG CGT CAG GAC	360
10	Gln Gln Gly Arg Ile Val Ala Leu Glu Gln Met Lys Arg Gln Asp	
11	510 515 520 525	
13	CAG GAG TGC CGA CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG	408
14	Gln Glu Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys	
15	530 535 540	
17	GGC ATA AAA AAG TTG ATC GGA AAT GTA CAG ATG CCA GTG GCT GCT GTA	456
18	Gly Ile Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val	
19	545 550 555	
21	GTT GTT ATG GCT TGC AAT CGG GCT GAT TAC CTG GAA AAG ACT ATT AAA	504
22	Val Val Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys	
23	560 565 570	
25	TCC ATC TTA AAA TAC CAA ATA TCT GTT GCG TCA AAA TAT CCT CTT TTC	552
26	Ser Ile Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe	
27	575 580 585	
29	ATA TCC CAG GAT GGA TCA CAT CCT GAT GTC AGG AAG CTT GCT TTG AGC	600
30	Ile Ser Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser	
31	590 595 600 605	
33	TAT GAT CAG CTG ACG TAT ATG CAG CAC TTG GAT TTT GAA CCT GTG CAT	648
34	Tyr Asp Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His	
35	610 615 620	
37	ACT GAA AGA CCA GGG GAG CTG ATT GCA TAC TAC AAA ATT GCA CGT CAT	696
38	Thr Glu Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His	
39	625 630 635	
41	TAC AAG TGG GCA TTG GAT CAG CTG TTT TAC AAG CAT AAT TTT AGC CGT	744
42	Tyr Lys Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg	
43	640 645 650	
45	GTT ATC ATA CTA GAA GAT GAT ATG GAA ATT GCC CCT GAT TTT TTT GAC	792
46	Val Ile Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp	
47	655 660 665	
49	TTT TTT GAG GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG	840
50	Phe Phe Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met	
51	670 675 680 685	
53	GCT ATT TCT TCT TGG AAT GAC AAT GGA CAA ATG CAG TTT GTC CAA GAT	888
54	Ala Ile Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp	
55	690 695 700	
57	CCT TAT GCT CTT TAC CGC TCA GAT TTT TTT CCC GGT CTT GGA TGG ATG	936
58	Pro Tyr Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met	
59	705 710 715	
61	CTT TCA AAA TCT ACT TGG GAC GAA TTA TCT CCA AAG TGG CCA AAG GCT	984
62	Leu Ser Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala	
63		

	720	725	730	
1	TAC TGG GAC GAC TGG CTA AGA CTC AAA GAG AAT CAC AGA GGT CGA CAA			
2	Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln			1032
3	735	740	745	
6	TTT ATT CGC CCA GAA GTT TGC AGA ACA TAT AAT TTT GGT GAG CAT GGT			
7	Phe Ile Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly			1080
8	750	755	760	765
9				
10	TCT AGT TTG GGG CAG TTT TTC AAG CAG TAT CTT GAG CCA ATT AAA CTA			
11	Ser Ser Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu			1128
12	770	775	780	
13				
14	AAT GAT GTC CAG GTT GAT TGG AAG TCA ATG GAC CTT AGT TAC CTT TTG			
15	Asn Asp Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu			1176
16	785	790	795	
17				
18	GAG GAC AAT TAC GTG AAA CAC TTT GGT GAC TTG GTT AAA AAG GCT AAG			
19	Glu Asp Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys			1224
20	800	805	810	
21				
22	CCC ATC CAT GGA GCT GAT GCT GTC TTG AAA GCA TTT AAC ATA GAT GGT			
23	Pro Ile His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly			1272
24	815	820	825	
25				
26	GAT GTG CGT ATT CAG TAC AGA GAT CAA CTA GAC TTT GAA AAT ATC GCA			
27	Asp Val Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala			1320
28	830	835	840	845
29				
30	CGG CAA TTT GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGT GCA			
31	Arg Gln Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala			1368
32	850	855	860	
33				
34	GCA TAT AAA GGA ATA GTA GTT TTC CGG TAC CAA ACG TCC AGA CGT GTA			
35	Ala Tyr Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val			1416
36	865	870	875	
37				
38	TTC CTT GTT GGC CAT GAT TCG CTT CAA CAA CTC GGA ATT GAA GAT ACT			
39	Phe Leu Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr			1464
40	880	885	890	
41				
42	TAA CAAAGATATG ATTGCAGGAG CCCGGGCAAATTTTGACT TATTGGGTAG			1517
43	*			
44				
45				
46	GATGCATCGA GCTGACACTA AACCATGATT TTACCAGTTA CATACAACGT TTTAATGTTA			1577
47				
48	TACGGAGGAG CTCACTGTT TAGTGTGAA GGGATATCGG CTTCTTAGTA TTGGATGAAT			1637
49				
50	CATCAACACA ACCTATTATT TTAAGTGTTC AGAACATAAA GAGGAAATGT AGCCCTGTAA			1697
51				
52	AGACTATACA TGGGACCATC ATAATCGCGG CCGCGAATT			1737
53				
54				
55				
56	(2) INFORMATION FOR SEQ ID NO: 4:			
57				
58	(i) SEQUENCE CHARACTERISTICS:			
59	(A) LENGTH: 446 amino acids			
60	(B) TYPE: amino acid			
61	(D) TOPOLOGY: linear			
62				
63	(ii) MOLECULE TYPE: protein			

1 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 2 Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu Ile Leu
 3 1 5 10 15
 4
 5 Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala Thr Gln
 6 20 25 30
 7
 8 Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His
 9 35 40 45
 10
 11 Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu Gln Gln
 12 50 55 60
 13
 14 Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp Gln Glu
 15 65 70 75 80
 16
 17 Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
 18 85 90 95
 19
 20 Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val Val Val
 21 100 105 110
 22
 23 Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys Ser Ile
 24 115 120 125
 25
 26 Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
 27 130 135 140
 28
 29 Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Asp
 30 145 150 155 160
 31
 32 Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His Thr Glu
 33 165 170 175
 34
 35 Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
 36 180 185 190
 37
 38 Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg Val Ile
 39 195 200 205
 40
 41 Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Phe Phe
 42 210 215 220
 43
 44 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
 45 225 230 235 240
 46
 47 Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp Pro Tyr
 48 245 250 255
 49
 50 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
 51 260 265 270
 52
 53 Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
 54 275 280 285
 55
 56 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
 57 290 295 300
 58
 59 Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
 60 305 310 315 320
 61
 62 Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
 63

1 325 330 335
 2 Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
 3 340 345 350
 4
 5 Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
 6 355 360 365
 7
 8 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
 9 370 375 380
 10
 11 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala Arg Gln
 12 385 390 395
 13
 14 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
 15 405 410 415
 16
 17 Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val Phe Leu
 18 420 425 430
 19
 20 Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr
 21 435 440 445
 22
 23
 24
 25 (2) INFORMATION FOR SEQ ID NO: 5:
 26
 27 (i) SEQUENCE CHARACTERISTICS:
 28 (A) LENGTH: 1854 base pairs
 29 (B) TYPE: Nucleotide
 30 (C) STRANDEDNESS: double
 31 (D) TOPOLOGY: linear
 32
 33 (ii) MOLECULE TYPE: cDNA to mRNA
 34
 35 (iii) HYPOTHETICAL: No
 36
 37 (iv) ANTI-SENSE: No
 38
 39 (vi) ORIGINAL SOURCE:
 40 (A) ORGANISM: Arabidopsis thaliana
 41 (B) STRAIN: Columbia
 42 (D) DEVELOPMENTAL STAGE: Mature plants
 43 (F) TISSUE TYPE: All tissues
 44
 45 (vii) IMMEDIATE SOURCE:
 46 (A) LIBRARY: Lambda Uni-ZAP (EcoRI/XbaI) and
 47 Lambda ACT (XbaI)
 48 (B) CLONE: pBSK-Ara-GntI-full #8
 49
 50 (ix) FEATURE:
 51 (A) NAME/KEY: misc_feature
 52 (B) LOCATION: 1185..1193
 53 (D) OTHER INFORMATION: /function= "Asn Codon is a
 54 potential glycosylation site"
 55 /product= "Consensus sequence for
 56 N-glycosylation"
 57 /phenotype= "N glycans modulate
 58 protein characteristics"
 59 /standard_name= "N glycosylation site"
 60 /label= pot-CHO
 61 /note= "absent in animal GntI sequences"
 62
 63 (ix) FEATURE:

	15	20	25	
1				
2	ACG CAA TCA CAG TAT GCA GAT CGC CTC AGT TCC GCT ATC GAA TCT GAG			266
3	Thr Gln Ser Gln Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu			
4	30	35	40	
5				
6	AAC CAT TGC ACT AGT CAA ATG CGA GGC CTC ATA GAT GAA GTT AGC ATC			314
7	Asn His Cys Thr Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile			
8	45	50	55	60
9				
10	AAA CAG TCG CGG ATT GTT GCC CTC GAA GAT ATG AAG AAC CGC CAG GAC			362
11	Lys Gln Ser Arg Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp			
12	65	70	75	
13				
14	GAA GAA CTT GTG CAG CTT AAG GAT CTA ATC CAG ACG TTT GAA AAA AAA			410
15	Glu Glu Leu Val Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys			
16	80	85	90	
17				
18	GGA ATA GCA AAA CTC ACT CAA GGT GGA CAG ATG CCT GTG GCT GCT GTA			458
19	Gly Ile Ala Lys Leu Thr Gln Gly Gly Gln Met Pro Val Ala Ala Val			
20	95	100	105	
21				
22	GTG GTT ATG GCC TGC AGT CGT GCA GAC TAT CTT GAA AGG ACT GTT AAA			506
23	Val Val Met Ala Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys			
24	110	115	120	
25				
26	TCA GTT TTA ACA TAT CAA ACT CCC GTT GCT TCA AAA TAT CCT CTA TTT			554
27	Ser Val Leu Thr Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe			
28	125	130	135	140
29				
30	ATA TCT CAG GAT GGA TCT GAT CAA GCT GTC AAG AGC AAG TCA TTG AGC			602
31	Ile Ser Gln Asp Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser			
32	145	150	155	
33				
34	TAT AAT CAA TTA ACA TAT ATG CAG CAC TTG GAT TTT GAA CCA GTG GTC			650
35	Tyr Asn Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val			
36	160	165	170	
37				
38	ACT GAA AGG CCT GGT GAA CTG ACT GCG TAC TAC AAG ATT GCA CGT CAC			698
39	Thr Glu Arg Pro Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His			
40	175	180	185	
41				
42	TAC AAG TGG GCA CTG GAC CAG TTG TTT TAC AAA CAC AAA TTT AGT CGA			746
43	Tyr Lys Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg			
44	190	195	200	
45				
46	GTG ATT ATA CTA GAA GAC GAT ATG GAA ATT GCT CCA GAC TTC TTT GAT			794
47	Val Ile Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp			
48	205	210	215	220
49				
50	TAC TTT GAG GCT GCA GCT AGT CTC ATG GAT AGG GAT AAA ACC ATT ATG			842
51	Tyr Phe Glu Ala Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met			
52	225	230	235	
53				
54	GCT GCT TCA TCA TGG AAT GAT AAT GGA CAG AAG CAG TTT GTG CAT GAT			890
55	Ala Ala Ser Ser Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp			
56	240	245	250	
57				
58	CCC TAT GCG CTA TAC CGA TCA GAT TTT TTT CCT GGC CTT GGG TGG ATG			938
59	Pro Tyr Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met			
60	255	260	265	
61				
62	CTC AAG AGA TCG ACT TGG GAT GAG TTA TCA CCA AAG TGG CCA AAG GCT			986
63				

1 Leu Lys Arg Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala
 2 270 275 280
 3
 4 TAC TGG GAT GAT TGG CTG AGA CTA AAG GAA AAC CAT AAA GGC CGC CAA 1034
 5 Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln
 6 285 290 295 300
 7
 8 TTC ATT GCA CCG GAA GTC TGT AGA ACA TAC AAT TTT GGT GAA CAT GGG 1082
 9 Phe Ile Ala Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly
 10 305 310 315
 11
 12 TCT AGT TTG GGA CAG TTT TTC AGT CAG TAT CTG GAA CCT ATA AAG CTA 1130
 13 Ser Ser Leu Gly Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu
 14 320 325 330
 15
 16 AAC GAT GTG ACG GTT GAC TGG AAA GCA AAG GAC CTG GGA TAC CTG ACA 1178
 17 Asn Asp Val Thr Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr
 18 335 340 345
 19
 20 GAG GGA AAC TAT ACC AAG TAC TTT TCT GGC TTA GTG AGA CAA GCA CGA 1226
 21 Glu Gly Asn Tyr Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg
 22 350 355 360
 23
 24 CCA ATT CAA GGT TCT GAC CTT GTC TTA AAG GCT CAA AAC ATA AAG GAT 1274
 25 Pro Ile Gln Gly Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp
 26 365 370 375 380
 27
 28 GAT GAT CGT ATC CGG TAT AAA GAC CAA GTA GAG TTT GAA CGC ATT GCA 1322
 29 Asp Asp Arg Ile Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala
 30 385 390 395
 31
 32 GGG GAA TTT GGT ATA TTT GAA GAA TGG AAG GAT GGT GTG CCA CGA ACA 1370
 33 Gly Glu Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr
 34 400 405 410
 35
 36 GCA TAT AAA GGA GTA GTG GTG TTT CGA ATC CAG ACA ACA AGA CGT GTA 1418
 37 Ala Tyr Lys Gly Val Val Val Phe Arg Ile Gln Thr Arg Arg Val
 38 415 420 425
 39
 40 TTC CTG GTT GGG CCA GAT TCT GTA ATG CAG CTT GGA ATT CGA AAT TCC 1466
 41 Phe Leu Val Gly Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser
 42 430 435 440
 43
 44 TGA TGC~~AAAACAT~~ ATGAAAGGAA AAGAACATTG TGGACCGCAT GCAGCCTCCT 1519
 45 *
 46 445
 47
 48 TCTAGCAGCT GTTAGGTTGT ATTGTTATTT ATGGATGAGT TTGTAGAGCG GTGGGGTTAA 1579
 49
 50 CTTAACAGC AAGGAAGCTC TGGTGACCAG GCTGATTGGC TTAGAACGTTA TGGGAACCCC 1639
 51
 52 TTGAAAGGGT CAGGGTTAAA TATATTCAG TTGTTTTATT AGTGATTATC TTGTGGGTAA 1699
 53
 54 CTTATACGAA TGCAAATCAT TCTATGCAGT TTTCTTCGT CCCACTTGTT TTGGCTTCTC 1759
 55
 56 TATTGCTAGT GTACATATCT CTTCAAACAT GTACTAAATA ATGCGTGTG CTTCAAAGAA 1819
 57
 58 GTAACTTTA TTAAAAAAA AAAAAAAAC TCGAG 1854
 59
 60
 61 (2) INFORMATION FOR SEQ ID NO: 6:
 62
 63 (i) SEQUENCE CHARACTERISTICS:

```

1 (A) LENGTH: 444 amino acids
2 (B) TYPE: Amino acid
3 (D) TOPOLOGY: Linear
4
5 (ii) MOLECULE TYPE: Protein
6 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
7
8 Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu Ile Pro Ala Ala
9 1 5 10 15
10 Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln Thr Gln Ser Gln
11 20 25 30
12
13 Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu Asn His Cys Thr
14 35 40 45
15
16 Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile Lys Gln Ser Arg
17 50 55 60
18
19 Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp Glu Glu Leu Val
20 65 70 75 80
21
22 Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys Gly Ile Ala Lys
23 85 90 95
24
25 Leu Thr Gln Gly Gly Gln Met Pro Val Ala Ala Val Val Val Met Ala
26 100 105 110
27
28 Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys Ser Val Leu Thr
29 115 120 125
30
31 Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp
32 130 135 140
33
34 Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser Tyr Asn Gln Leu
35 145 150 155 160
36
37 Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val Thr Glu Arg Pro
38 165 170 175
39
40 Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala
41 180 185 190
42
43 Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg Val Ile Ile Leu
44 195 200 205
45
46 Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Tyr Phe Glu Ala
47 210 215 220
48
49 Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met Ala Ala Ser Ser
50 225 230 235 240
51
52 Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp Pro Tyr Ala Leu
53 245 250 255
54
55 Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Lys Arg Ser
56 260 265 270
57
58 Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp Asp
59 275 280 285
60
61 Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln Phe Ile Ala Pro
62 290 295 300
63

```

1 Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu Gly
2 305 310 315 320
3
4 Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val Thr
5 325 330 335
6
7 Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr Glu Gly Asn Tyr
8 340 345 350
9
10 Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg Pro Ile Gln Gly
11 355 360 365
12
13 Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp Asp Asp Arg Ile
14 370 375 380
15
16 Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala Gly Glu Phe Gly
17 385 390 395 400
18
19 Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr Ala Tyr Lys Gly
20 405 410 415
21
22 Val Val Val Phe Arg Ile Gln Thr Thr Arg Arg Val Phe Leu Val Gly
23 420 425 430
24
25 Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser
26 435 440
27
28